

# 冯裕祺

(+86) 182-7839-3372 | 853346463@qq.com|[个人主页](#)  
求职意向：算法工程师,随时到岗，实习时长三个月到六个月



## 教育经历

东北大学 (985/211) (免试保送)	应用统计	硕士	2020.09-至今 (2023 年 6 月毕业)
获得荣誉：研究生一等奖学金，华为杯数学建模参赛奖等。			
桂林理工大学	应用统计	本科	2016.09-2020.06
学习成绩：GPA(93/100)英语水平：CET-4 532 CET-6 489,能够与外国客户简单商务交流			
获得荣誉：挑战杯全国三等奖、挑战杯省级特等奖、金种子奖学金奖等。			

## 项目经历

Mintech, ToB 部门风控模型实习生		2022.3.9-至今
<ul style="list-style-type: none"><li>● <b>岗位描述</b>：负责与巴基斯坦某银行的个人信贷业务风控模型建立以及信贷额度策略调整。</li><li>● <b>工作内容</b>：从合作方提供的 mongoDB 数据库中提取用户数据，输出监控报告，进行 vintage 分析和用户画像构建。根据历史用户数据，清洗数据并构建特征，针对逾期用户与为逾期用户类不平衡问题使用采样和加权方法进行处理。</li><li>● <b>使用方法</b>：通过使用欠采样方法结合自适应加权并结合 xgb,logistic regression 和随机森林等方法进行集成学习，最终提升模型识别率 ks 值 5%。</li></ul>		
超高维小样本分类问题	论文	2021.06-2021.10
<ul style="list-style-type: none"><li>● <b>项目描述</b>：对超高维和小样本的分类问题，提出了使用 ball correlation(相比皮尔逊相关系数更好的方法)结合 abess(最优子集切片算法)方法的分类方法。</li><li>● <b>方法</b>：通过两阶段方法首先使用 ball 方法对样本进行初步的特征筛选，然后将筛选出的特征使用 abess 方法进行分类。</li><li>● <b>结果</b>：提出的改进模型在 microarray 数据集（DNA 基因序列数据，用于癌症判别诊断）上面的表现经过交叉验证等方法得到的结果明显优于传统的支持向量机，逻辑回归和 knn 等方法，分类 ACC 提升 15%左右，计算时间极大缩短。</li></ul>		
统计变量选择方法汇总 (超链接)	作品	2020.09-2021.09
<ul style="list-style-type: none"><li>● <b>作品描述</b>：汇总整理统计中传统的收缩方法 LASSO(L1 范数惩罚), ridge(L2 范数惩罚), SCAD 和 elastic net 等经典收缩方法从原文中总结其原理和特点。并进一步总结出前沿的 screening 方法（针对超高维问题方法）并且将是用统计变量选择方法的应用论文进行汇总。</li><li>● <b>总结</b>：通过对统计变量选择方法按照时间顺序汇总整理，对整个方法发展的历程和思路有了较为清晰的认识。并且通过阅读原文将其原理深度了解，对在实际问题中解决变量选择问题有很好的帮助。</li></ul>		
呼吸机智能辅助脱机系统建模	东软汉枫医疗科技有限公司	2020.11-2020.12
<ul style="list-style-type: none"><li>● <b>项目描述</b>：为了评估 ICU 患者是否撤离呼吸机，基于 MIMIC III 数据库，结合盛京医院搜集到的数据，建立呼吸机撤机预报模型，减少 ICU 中医护工作量，提高撤机准确率。</li><li>● <b>模型建立</b>：利用 Python 语言使用多种算法建立模型，最终利用 Voting 的方法将随机森林、XGBoost 以及多层感知机算法进行集成，在测试集上达到 85%的准确率。</li><li>● <b>项目成果</b>：该方案已在沈阳市盛京医院开展初步临床对照实验，准确率达到 80%，极大减少了医护的工作量，未来将应用于 ICU 呼吸重症病房。</li></ul>		

## 专业技能

- **具备技能**：掌握 python,R,pytorch,sklearn,SQL,mongoDB,linux 系统。
- **个人作品**：[kaggle 房价预测（全球排名 700 / 15000）](#)，[高维类不平衡问题自适应加权方法](#)